

国外电子信息科学教材系列

# 多媒体数字压缩 原理与标准

Digital Compression for Multimedia  
**PRINCIPLES & STANDARDS**

Jerry D.Gibson

Toby Berger

[美] Tom Lookabaugh 著

Dave Lindbergh

Richard L.Baker

李煜晖 朱山风 段上为 等译

李 臻 审校



电子工业出版社

PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

URL: <http://www.phei.com.cn>

国外电子信息科学教材系列

# 多媒体数字压缩 原理与标准

Digital Compression for Multimedia  
Principles & Standards

[美] Jerry D. Gibson Toby Berger Tom Lookabaugh 著  
Dave Lindbergh Richard L. Baker

李煜晖 朱山风 段上为 等译

李 臻 审校

电子工业出版社

Publishing House of Electronics Industry

711001

北京·BEIJING

..

## 内 容 提 要

本书的作者均为信息理论方面的杰出专家,有 IEEE 信息理论分会主席、技术公司的总裁及技术标准的制定者,他们的携手合作使得本书在理论深度和技术实用性两方面都得到了兼顾。本书共 11 章,分别对数字压缩领域的背景知识、理论基础、算法以及技术标准作了全面介绍,在内容上深入浅出,在语言上通俗易懂,是关于数字压缩领域的一本优秀论著。

Original edition Copyright © 1998 by Morgan Kaufmann Publishers, Inc.

All rights reserved.

Chinese translation edition Copyright © 1999 by Publishing House of Electronics Industry.

All rights reserved.

本书中文简体专有翻译出版权由美国 Morgan Kaufmann Publishers, Inc. 授予电子工业出版社。该专有出版权受法律保护。

## 图书在版编目(CIP)数据

多媒体数字压缩原理与标准/(美)吉布森(Gibson, J. D.)等著;李煜晖等译. - 北京:电子工业出版社, 2000.8

国外电子信息科学教材系列

书名原文: Digital Compression for Multimedia Principles & Standards

ISBN 7-5053-5532-5

I. 多… II. ①吉…②李… III. 数字压缩IV. TP274

中国版本图书馆 CIP 数据核字(1999)第 38703 号

丛 书 名: 国外电子信息科学教材系列

书 名: 多媒体数字压缩原理与标准

著 者: [美] Jerry D. Gibson Toby Berger Tom Lookabaugh Dave Lindbergh Richard L. Baker

译 者: 李煜晖 朱山风 段上为 等

审 校 者: 李 臻

责任编辑: 刘文玲

特约编辑: 初登平

排版制作: 电子工业出版社计算机排版室

印 刷 者: 北京大中印刷厂

装 订 者: 三河市双峰装订厂

出版发行: 电子工业出版社 URL: <http://www.phei.com.cn>

北京市海淀区万寿路 173 信箱 邮编 100036

经 销: 各地新华书店

开 本: 787 × 1092 1/16 印张: 19.5 字数: 499.2 千字

版 次: 2000 年 8 月第 1 版 2000 年 8 月第 1 次印刷.

书 号: ISBN 7-5053-5532-5  
TP·2796

印 数: 4000 册 定价: 32.00 元

版权贸易合同登记号 图字:01-1999-1430

凡购买电子工业出版社的图书,如有缺页、倒页、脱页、所附磁盘或光盘有问题者,请向购买书店调换;若书店售缺,请与本社发行部联系调换。电话 68279077

## 出版说明

电子信息科学与技术的发展令人目不暇接,它正以主力军的姿态站在知识经济的最前列,引导着各行各业向着新的技术领域进军,影响着人们的工作和生活朝着现代化迈进。在电子信息科技领域里,除了计算机这一热门话题外,还有大量熟悉的、天天都会接触到的信息,如通信、网络、雷达、广播电视、自动控制、电磁干扰、信息安全……等等。在每一个专业领域,新技术、新产品都层出不穷。日新月异的科技发展,强烈地吸引着各方面的专家学者,及一切从事专业技术研究、生产、教学、管理、使用、维护的工程技术人员,他们执着地从国内外的技术文献和图书刊物里搜寻自己感兴趣的内容,不断地汲取新知识,增进才干、创造业绩。特别是正在从事教育或学习的院校师生,非常需要了解国外电子信息科学教育方面的发展,以使教与学更好地面向四个现代化,面向世界。为了满足读者的要求,我们电子工业出版社推出了这套“国外电子信息科学教材系列”丛书。

这套丛书的遴选是严格的,首先是选择了美国 Prentice Hall 出版公司等在世界很有影响的高校教材出版商,他们几十年来致力于教育图书的出版,其电子信息科技类教材在美国为众多大学所采用,并拥有不少各专业的资深作者和经典名著,在世界上产生了很大的影响。其次,我们把自己策划和准备引进选题的英文原版,请国内一些著名高等院校的学者教授予以评论挑选,有些还先期影印发行,以便广泛征求意见。在此基础上,我们约请北京地区的清华大学、北京大学、北京邮电大学、北京航空航天大学,上海地区的上海交通大学、复旦大学,南京地区的南京大学、解放军通信工程学院等名牌院校中在教学科研第一线的教师和博士(他们中间都是博士生导师和各专业的学科带头人,也有年轻的教师和博士)对原稿认真进行翻译加工,对其中一些表述不清甚至出错的地方也作了慎重地推敲或更正。这套丛书能高质量地与读者见面,应该感谢各位译者和审校者的辛勤劳动。

像任何事情一样,我们这项工作也会有不尽人意的地方。所以,在这套丛书出版之际,特别希望广大读者能给予指正、建议。我们将会陆续引进和出版新书,充实到这套丛书中来,以飨读者。

## 译 者 序

随着科技的飞速发展,多媒体、网络通信已成为我们生活中的一部分,人们正在进入数字化时代。作为数字化技术核心之一的数字压缩更是备受关注,本书对这个领域的背景知识、理论基础、算法以及技术标准作了全面的介绍,在内容上深入浅出,在语言上通俗易懂,是关于数字压缩领域的一本优秀论著。

本书的作者均为信息理论方面的杰出专家,有 IEEE 信息理论分会主席,也有技术公司的总裁,还有技术标准的制定者,他们的携手合作使得本书在理论深度和技术实用性两方面都得到了兼顾,对于通信技术和计算机技术的爱好者来说,这是一本易于入门的教材和读物,对于专业人员来说,它也有较高的参考价值。

本书共 11 章,李煜晖同志翻译了前 6 章及序言、附录,第 7、8、9 章由段上为翻译,朱山风翻译了第 10 章,姜东胜翻译了第 11 章。全书由李煜晖统稿,李旭晖、李克清、周敬华、韩光涛等同志也参与了翻译和资料查找工作,在翻译过程中,还得到沈祥兴、杨志峰等同志的大力协助,清华大学的李臻同志审校了全书,在此表示衷心感谢。

由于译者水平有限,经验不足,错误之处在所难免,望读者批评指正。

# 序 言

## 概述

本书论述了数字压缩方面的理论方法,发展趋势和压缩标准。在理论方法的覆盖面上涉及了从试探性压缩方法到基于信息论的严密方法。试探性方法以信号源的建模、信号处理技术和人们的“直觉”知识为基础。严密的信源编码技术则以信息编码和率失真理论作基础,这些都是信息论中处理数据有效压缩的分支。

Huffman 编码和变换编码的标量比特分配方法是最早运用于实际数字压缩系统的信息论方法,主要压缩计算机文件、声音、静态图像和视频图像等。在 20 世纪 60 年代和 70 年代,标量化取得了突破性进展,时域波形预测码、子带码、变换码相继出现。另外,发展了相关的感觉效果、信号源建模、信道差错机制以及编码性能评估。从 70 年代后期,这些和另外一些来源于信息论的严密方法越来越多地运用于数字压缩的实际系统当中。这当中关键性进展就是通用可逆编码技术的引进,特别是 Lempel-Ziv 理论和算术编码,以及率失真理论方法的应用,后者被用于设计和评估量化器和树形码,还用于设计解决约束系统的优化问题。今天,试探性方法和信息论方法已经融为一体,这种结合为信源数字压缩开辟了一条崭新的道路。

信号处理需要高性能的数字压缩方法。因此,数字压缩的进展与理论上的有关进展和信号处理技术的进展有着密切关系,例如:信号处理的研究人员可采用多维信号表示法、自适应算法和时域-频域表示法等理论工具。

许多先进的信号处理技术对计算机提出了很高的要求,因此,实现这些高性能技术就要有微型化领域和微电子领域不断有突破性进展,使得在计算机电路尺寸、复杂性和能源消耗上取得飞跃。今天的专用集成电路(ASICs),数字信号处理芯片(DSPs)和微处理器(CPUs)具有难以置信的能力,不可否认,这种进步还会继续下去。

另一个发展趋势是社会方面的,即知识爆炸——从工业化社会向信息化社会的转变。因特网,特别是 WWW 服务器的普及,企业网的流行,ISDN 的来临,ADSL 的发展前景以及调制解调器,无线通信和无线计算已经彻底地改变了世界,全球虚拟社会不再是虚拟的——它确实存在。

还有个发展趋势极大地促进了数字压缩技术的实用化,这就是产品标准化和通信、网络服务的兴起。目前,建立标准化的行动正以指数形式增长,这在各种通信领域,网络和压缩方法之间起到了协调作用,并为新产品和服务提供了良好的环境。

## 本书的创作

本书的创作可追溯到本文两位作者的一次对话,那是在 1990 年 IEEE 在圣地亚哥召开的信息论国际研讨会,Toby Berger 和 Jerry Gibson 发现,同以往的信源编码方面研讨会不同,这次

研讨会引起了一些小型研究团体的关注,他们热烈地讨论着圣地亚哥信源编码会议,并且,与会者除了一些理论专家以外还有许多来自企业的实际工作者。考虑到原来已经发表的方法和理论已趋于成熟,Toby 和 Jerry 提议举办一个短期培训,数字压缩的理论、实践及标准合为一体。为工程师、技术管理人员和那些想在多媒体应用领域建立、开发数字压缩系统的计算机科学家进行讲授。他的远景规划是,在举办几期培训之后,把讲义整理成书。UCLA 发展部赞同这个提议,并提供给了短训班更多的关于图像压缩特别是视频图像压缩的资料。为了达到预定目标,Gibson 又请了两位来自企业的专家——Tom Lookabaugh 和 Rich Baker,这两位长期从事视频协议、高性能视频压缩和标准制定方面的工作。从 1992 年起,Berger、Gibson、Lookabaugh 和 Baker 四人小组在 UCLA 举办了十二次多媒体数字压缩短训班,在 Compaq 举办了一次,在 NSA 举办了一次,在摩托罗拉大学举办了三次。创作这本书始于 1995 年。Dave Lindbergh,这位在视频协议标准化方面学识广博的专家,也于 1996 年加入了创作队伍,成为第五位作者。

## 本书的读者

本书适于在商业应用的数字压缩系统方面进行设计、制作和补充修改工作的工程师和科学家。也包括芯片设计人员、软件开发人员、通信系统工程师、电脑网络设计人员和信号处理工程师。同时我们希望这本书能对在这个领域从事领导研究的工作人员有所帮助,另外,本书可作为大学高年级学生以及在电子工程及计算机科学方面从事数字压缩工作的人员的课本或参考书。

第 1 章概述了本书的结构和内容,各章均可依各人的兴趣和背景知识而任意选读。但我们还是奉劝各位先读一下第 1 章,以便解除数字压缩领域的概况及其起源,同时它也概述了本书的结构和重点。有的读者可能愿意直接阅读有关压缩标准的章节,另一些愿意多了解一下背景知识的读者可在前面几章参考一些预备材料。

## 方法

为了让这本书易于理解,又能达到适当的精确性,我们在必要的地方涉及了数学基础和算法细节,但是,我们尽可能地简化这些内容,并把它们放在附录或专门的段落里,算法的细节很重要,从我们在短训班的经验来看,短训班的学生在没有透彻理解基本规律时,是无法掌握实际工作中需要的东西。另一方面,读者也将看到,这本书并不是一本经典学术著作式的教科书,我想,它恰到好处地综合了指导、标准的发展以及应用等方面的知识。

## 致谢

在此,对那些帮助我们搜集资料和出版此书的人,表达我们的谢意。UCLA 发展部的 Bill Goodin,他为我们提供机会来培训数字压缩短训班。我们也缘于此才结识了许多天才的工程师、科学家和其他一些参加培训并在工作中用到这些资料的专家。我们还要感谢 Morgan Kaufmann 出版社的 Elisabeth Beller 和 Jennifer Mann。主编 Elisabeth 出色的工作使得一切按计划进行并达到最高的专业水准。我们的编辑 Jennifer Mann 不厌其烦地建议、协调、敦促我们完成

此书。没有她的贡献就不会有那么多出色的评论家来批评、讨论和评点本书。这些评论家有 Barry Haskell(朗讯公司), William Pearlman (RPI), Bob Holm (Intel), Bob Safranek (AT&T), Bruce Walk (IBM), Ephraim Feig (IBM), Eve Riskins(华盛顿大学), John Limb(乔治亚技术学院), Mark Perkins(算法和应用科学杂志)。Michael Marcellin (Arizona 大学)。Ping Wah Wong (HP 实验室), Richard Cox (AT&T 实验室), Roger Dressler (杜比实验室), Roy Hoffman (IBM) and Tim Midwinter (British Telecom), 他们为我们提供了有价值的见解, 我们还要特别提一下 Rich Cox, 他非常认真, 花费了数小时仔细阅读了手稿并就这个专题坦率地发表了意见, 令我们十分感动。我们还要感谢我们的同事和学生, 他们提出一些见解并发现了一些印刷错误。后面附的名单只是其中的一部分, 我们很抱歉遗漏了相当多的人。

• Toby Berger 特别鸣谢 Reymond Yeung、Ram Zamir、Victor wei、Srikant Jayaraman、Harish Viswanathan、P. Subrahmanya、Ashok Montravadi、James Chow、Adi Wyner、Ed Mosteig、Luis Lastros、康奈尔大学, EE562 和 EE567, 费吉尼亚大学 EE717。

• Jerry Gibson 鸣谢 Madhukar Budagavi、Hui Dong、Mark Kokes、Justin Ridge、Roderick Maddox、Fred Ware、Wenye Yang、Mark Randolph、sharif sazzard、Jason Brewer、Vince Rhee、Stan McCellan、Myron Moodie、Seung Nam、Insung Lee、Yoon Chae Cheong、Hong Chae Woo、Dae Gwon Jeong、Victor Taylor、Boneung Koo、Wen Whei Chang、Steven Gray、Khalid Sayood、David Comstock、Jane Asmuth、Richard Fenwick、Charles Moore、David Purcell、John McVay、Andy Goris、Louis Sauter、Victor Berglund、和 Ed Cross。

• Tom Lookabaugh 鸣谢 Matt Goldman、Mike Perkins、Didier Le Gall、Bill Helms、Christie Caldwell 和 Robin Wilson。

• Dave Lindbergh 鸣谢 Jeff Bernstein、Suneel Bhagat、Steve Botzko、Tony Crossman、Mike Nilsson、Sakae Okubo、Mark Reid、Gary Sullivan、John Villasenor、Hong Wang 和 Bob Webber。

• 我们共同致谢 UCLA 短训班的所有学员。

数字压缩是一个令人兴奋而且重要的领域, 数字压缩的应用与日俱增, 我们希望这本书能给你带来乐趣、知识和收益。

## 作者介绍

Jerry D. Gibson 现任德克萨斯州 Southern Methodist 大学电子工程系主任, 他同时还在 Notre Dane 大学, Nebraska - Lincoln 大学和德克萨斯州 A&M 大学兼任教席。1996 年曾出任 IEEE 信息理论分会主席。

Gibson 博士是《移动通信手册》(1995)和《通信大全》(1997)两书的主要编者, 1990 年 Gibson 博士荣获美国工程教育协会的 Fredrick Emmons Terman 奖金, 1992 年因其在自适应预测和语音编码的理论和实践方面作出的贡献被选为 IEEE 会员。1993 年因其在语音处理领域方面的成就与他人共同获得了 IEEE 信号处理分会的 Sensor Paper 奖。

Gibson 博士的研究领域包括数据, 语音, 图像和视频压缩, 网络多媒体, 无线通信, 信息论和数字信号处理。



Toby Berger 于 1940 年 9 月 4 日生于纽约。1962 年获得耶鲁大学的电子工程学士学位。分别在 1964 年和 1966 年在马萨诸塞州的哈佛大学获得了应用数学的硕士和博士学位。

1962 ~ 1968 年间,他作为麻省 Raytheon 公司的首席研究专家,致力于通信理论信息论和一致信号处理方面的研究。1968 年他进入了纽约州康乃尔大学,现任工科的 J. Preston Levis 教授,他的研究领域包括多终端编码理论,随机领域内的信息论,通信网络视频压缩,手写体压缩以及一致性信号处理。他还编写了教科书:《率失真理论——数字压缩的数学基础》。

Berger 教授还是《IEEE 信息理论汇编》的主编和 IEEE 信息论分会的主席,他曾担任过 Guggenheim 基金会,日本科学普及协会,中华人民共和国教育部外籍专家组以及 Fulbright 基金会的成员,1982 年他因为在“电子工程教育方面作出的突出贡献”荣获美国电子教育协会的 Frederick E. Terman 奖金。Berger 博士现为 IEEE 会员。

Tom Lookabaugh 是 DiviCom 公司总裁,作为公司的创始人之一,Lookabaugh 教授将其视频,音频压缩与高性能复用系统方面的研究成果和他在处理与发展合作关系方面的经验一同带入了 DiviCom 公司。

在加盟 DiviCom 公司以前,Lookabaugh 教授在 CLI 公司干了 5 年,成为研发与新产业技术方面的执行董事,他在 CLI 工作期间,参与了多个压缩系统的算法开发工作,其中包括视频会议与广播视频的三代产品。在与产品开发和测试方面的用户合作中,他与之建立了良好的关系。他承担的任务中有一项就是视频点播系统的 MPEG - 1 解码器的开发。

Lookabaugh 在科罗拉多大学获得工程物理学学士学位;在斯坦福大学获得电子技术,统计和工程管理硕士学位,并获得电子工程博士学位。

Dave Lindbergh 从 1997 年开始就从事计算机和数据通信方面的工作,1981 年他开发了 Lindbergh 系统,这是一个微机数据通信软件的开发平台,其中包括获奖的 OMNITERM 软件。作为一个资深工程师,他设计了调制解调器协议与软件并开发了现在广泛应用于调制解调器领域的 APT(异步性能测试器)数据通信衡量工具。

在 1990 年到 1992 年间担任一家多媒体公司——CD Atlas 公司的总裁。

1993 年他加入了 PictureTel 集团,这是世界领先的视频会议设备制造商,现在他已是 PictureTel 技术标准部的经理了。在美国和国际标准化组织包括 ANSI T1 和 TIA 以及 ITU 中,他是 PictureTel 公司的代表,是 ITU - T H.223, H.224 和 H.281 标准的主要制定者,也是 ITU H.324 的制定者,现任 ITU H.324 系统专家组的主席。他是数字压缩技术的一项美国专利的持有者,IEEE 的会员。

Richard L. Baker 博士,马萨诸塞州丹佛市 PictureTel 公司的副总裁和首席技术官员,从事桌面可视通信方面的关键技术开发。在 1990 年 1 月加入公司以前,他是 UCLA 的电子工程助理教授,在那里他讲授信息论,率失真理论和电路分析方面的课程,并从事图像和视频压缩技术的研究。Baker 博士已出版过 30 多本技术书籍。

Baker 教授还是国际多媒体远程会议组织(IMTC)的董事会成员和副总裁,IMTC 是一家非盈利性组织,其成员超过 100 家公司,主要致力于基于 ITU 的 H.32X/T.120 系列标准的产品开发。在 IMTC 中,他还主持了网络包会议(Packet Network Conferencing)研究小组的工作。

# 目 录

第 1 章 数字压缩引论	(1)
1.1 为什么要压缩	(1)
1.2 数字压缩问题	(2)
1.2.1 数字压缩的同义词	(2)
1.2.2 数字压缩问题的组成	(3)
1.2.3 数字压缩问题的类型	(4)
1.3 输入信源的格式	(4)
1.4 信源重构的品质	(5)
1.4.1 性能衡量	(6)
1.4.2 失真度的主观评价	(6)
1.5 系统要素和性能比较	(7)
1.6 应用和标准	(8)
1.7 本书概述	(9)
第 2 章 无损信源编码	(11)
2.1 引言	(11)
2.2 瞬时可变长码	(11)
2.3 可唯一译码性(UD)	(14)
2.4 Huffman 编码	(15)
2.5 非二进制 Huffman 编码	(17)
2.6 Kraft 不等式及最优化	(18)
2.7 3类和 4类传真标准	(19)
2.7.1 3类传真	(19)
2.7.2 4类传真	(25)
2.7.3 噪声和二值图像(Half-Toning)	(27)
2.8 画线压缩	(27)
2.9 熵和性能界	(31)
2.9.1 一些不等式	(31)
2.9.2 熵	(31)
2.9.3 可达到熵下界的压缩	(33)
2.10 条件熵和互信息	(35)
2.11 平稳信源的熵率	(36)
2.11.1 联合熵和链规则	(37)
2.11.2 熵率的定义	(38)
2.11.3 Shannon-Fano 码	(40)

<b>第 3 章 通用无损信源编码</b> .....	(41)
3.1 自适应性和通用性 .....	(41)
3.2 划分 .....	(42)
3.3 LZ 压缩 .....	(44)
3.3.1 LZ78 .....	(44)
3.3.2 LZW .....	(47)
3.3.3 LZY .....	(50)
3.3.4 LZ77 .....	(52)
3.3.5 LZ77、LZW 和 LZY 的性能比较 .....	(55)
3.3.6 冗余度压缩.....	(58)
3.4 Elias 编码、算术编码和 JBIG 传真 .....	(60)
3.4.1 Elias 编码 .....	(61)
3.4.2 算术编码.....	(64)
3.4.3 JBIG 传真标准 .....	(65)
<b>第 4 章 量化</b> .....	(72)
4.1 引言 .....	(72)
4.2 标量量化.....	(72)
4.2.1 均匀量化.....	(72)
4.2.2 非均匀量化.....	(75)
4.2.3 对数压扩.....	(77)
4.2.4 自适应量化.....	(78)
4.2.5 嵌入量化.....	(80)
4.3 矢量量化.....	(82)
4.3.1 VQ 结构、设计和性能.....	(82)
4.3.2 最优化的 VQ .....	(84)
4.3.3 结构化的 VQ .....	(84)
4.4 小结 .....	(88)
<b>第 5 章 预测编码</b> .....	(89)
5.1 引言 .....	(89)
5.2 线性预测模型和线性预测编码 .....	(90)
5.2.1 系数的计算.....	(91)
5.2.2 其他参数.....	(94)
5.2.3 浊音/清音判定和激励信号 .....	(95)
5.2.4 基音周期估计.....	(96)
5.2.5 激励增益.....	(97)
5.2.6 LPC 性能 .....	(98)
5.3 增量调制和差分 PCM .....	(98)
5.3.1 增量调制 .....	(100)
5.3.2 Nyquist 采样预测编码 .....	(102)
5.3.3 短时预测器自适应 .....	(104)
5.4 嵌入式 DPCM .....	(106)

5.5	多脉冲线性预测编码(MPLPC)	(107)
5.6	码激励线性预测编码	(110)
5.7	感知权重和后滤波	(113)
5.8	小结	(117)
<b>第 6 章</b>	<b>线性预测语音编码标准</b>	<b>(118)</b>
6.1	引言	(118)
6.2	ITU G.721/G.726/G.727	(119)
6.3	美国联邦标准 1015	(121)
6.4	美国联邦标准 1016	(124)
6.5	GSM 13 kbps 编码器	(125)
6.6	TIA 8kbps VSELP	(127)
6.7	TIA QCELP	(131)
6.8	LD - CELP, ITU G.728	(133)
6.9	ITU G.729	(135)
6.10	ITU G 723.1	(138)
6.11	JDC(PDC)全速率、GSM 半速率和 JDC 半速率标准	(140)
6.12	2.4kbps 的美国联邦标准	(141)
6.13	其他的标准和即将执行的标准	(144)
<b>第 7 章</b>	<b>频域编码</b>	<b>(145)</b>
7.1	引言	(145)
7.2	语音子带编码	(145)
7.3	图像的子带编码	(150)
7.4	声音和图像的变换编码	(155)
7.4.1	离散变换	(156)
7.5	小波编码	(163)
7.6	分形编码	(166)
7.7	小结	(168)
<b>第 8 章</b>	<b>频域语言和音频编写标准</b>	<b>(169)</b>
8.1	引言	(169)
8.2	ITU G.722 宽带音频和低码率扩充	(169)
8.3	音频的同时屏蔽和暂时屏蔽	(172)
8.4	为视频设计的高质量音频标准	(174)
8.4.1	MPEG - 1 音频	(175)
8.4.2	MPEG - 2 音频	(176)
8.4.3	Dolby AC - 2 和 AC - 3	(179)
8.4.4	AT&T 感知音频编码器	(182)
8.5	音频存储设备编码	(183)
8.5.1	DCC PASC 编码器	(183)
8.5.2	微型磁盘 ATRAC 编码器	(184)

8.6	国际海事卫星语音编码器	(184)
8.7	小结	(185)
<b>第9章</b>	<b>JPEG 静止图像压缩标准</b>	<b>(186)</b>
9.1	引言	(186)
9.2	基本 JPEG	(186)
9.3	累进编码	(189)
9.4	锥形(金字塔)编码	(189)
9.5	熵编码	(192)
9.5.1	DCT 系数编码举例	(194)
9.6	图像数据的规定	(195)
9.7	无损编码模式	(196)
9.8	小结	(197)
<b>第10章</b>	<b>多媒体会议标准</b>	<b>(199)</b>
10.1	引言	(199)
10.2	基于 ISDN 的 H.320 视频会议	(200)
10.2.1	H.320 标准簇	(200)
10.2.2	H.221 复用	(201)
10.2.3	系统控制协议	(204)
10.2.4	音频编码	(205)
10.2.5	视频编码	(207)
10.2.6	H.231 和 H.243——多点会议	(208)
10.2.7	H.233 和 H.234——加密	(209)
10.2.8	H.224 和 H.281——实时远端摄像控制	(210)
10.2.9	H.331 广播	(210)
10.3	H.320 网络适配标准 H.321 和 H.322	(210)
10.3.1	H.321 - H.320 适配于 ATM 和 B-ISDN	(211)
10.3.2	H.322 - H.320 适配于 ISO 以太网	(211)
10.4	新一代标准 H.323、H.324 和 H.310	(211)
10.4.1	H.245 控制协议	(212)
10.4.2	音频和视频编解码	(213)
10.4.3	用于分组交换网络的 H.323	(215)
10.4.4	用于低比特率电路交换网络的 H.324	(219)
10.4.5	用于 ATM 和 B-ISDN 网络的 H.310	(225)
10.5	用于数据会议和会议控制的 T.120	(225)
10.5.1	T.120 基础结构	(226)
10.5.2	T.120 应用协议	(227)
10.6	多媒体会议系统中的延迟	(228)
10.6.1	音频延迟来源	(228)
10.7	小结	(231)
<b>第11章</b>	<b>MPEG 压缩</b>	<b>(232)</b>
11.1	引言	(232)

11.2	MPEG 模型	(232)
11.2.1	主要应用和问题	(232)
11.2.2	制定标准的策略	(234)
11.2.3	部分 MPEG-1 和 MPEG-2 标准	(237)
11.3	MPEG 视频	(237)
11.3.1	基本算法	(237)
11.3.2	短暂预测	(239)
11.3.3	频域分解	(245)
11.3.4	量化	(246)
11.3.5	可变长编码	(247)
11.3.6	MPEG 语法分层	(250)
11.3.7	比特率控制	(250)
11.3.8	限制参数、层次和轮廓	(252)
11.4	MPEG 音频	(253)
11.4.1	层次	(253)
11.4.2	基本算法	(253)
11.4.3	子带分解	(254)
11.4.4	定标、量化和编码	(254)
11.4.5	多通道压缩	(255)
11.5	MPEG 系统	(256)
11.5.1	定时	(256)
11.5.2	系统和程序流	(256)
11.5.3	传输流	(258)
11.5.4	打包基本流(PES)和 MPEG 包	(258)
11.5.5	节目指定信息	(259)
11.6	MPEG 扩展	(260)
11.6.1	MPEG-4	(260)
11.6.2	数字存储媒体命令和控制	(261)
11.6.3	高级音频编码(AAC)	(262)
11.6.4	专业化或 4:2:2 纲要	(262)
11.6.5	MPEG-7	(262)
11.7	小结	(262)
附录A	语音质量及清晰度	(264)
A.1	引言	(264)
A.2	语音编码器评估的阶段	(264)
A.3	非正式测试	(264)
A.3.1	客观测量	(265)
A.3.2	主观测试	(265)
A.4	正式测试	(266)
A.4.1	清晰度	(266)
A.4.2	质量	(266)

A.5 重要因素 .....	(268)
附录 B Huffman 编码使 $\bar{l}$ 极小化的证明 .....	(269)
附录 C 每一种 UD 编码均满足 Kraft 不等式的证明 .....	(271)
附录 D 趋近熵率的方式 .....	(273)
附录 E LZJ 前向特性的证明 .....	(275)
附录 F LZ77 的 $L_k$ 有效编码 .....	(276)
词汇表 .....	(277)
参考文献 .....	(287)

# 第 1 章 数字压缩引论

## 1.1 为什么要压缩

娱乐、电信、因特网——这些都是我们日常生活的一部分，在生活中我们享受它们，使用它们，在报纸和杂志上读到它们，在电视上看到它们，并在这些东西上花钱，但是我们已经拥有收音机、电视机、音响、电话几十年了。现在有什么新东西呢？这些东西不是早就已经存在了吗？——不对，至少不全对！这里有个新词，就是“数字化”。今天我们到处谈论着数字通信系统和网络，还有数字化的电影、电视、音乐、图像和声音。为什么要数字化呢？因为数字信号易于存储和远距离传输，而且没有累积失真，数字化存储的信息(如：音乐)可高品质地被还原。

但是也有不利的一面，一些重要的信号，像声音、音乐、电视、电影的数字化版本需要每秒更多地比特数去存储或传输，这样就造成了高成本。例如，表 1.1 (Jayaut, Johnston, and Safranek 1993)列出了几种重要信源信号的原始数据速率(未经压缩)。当然，许多数据看起来很大，但这些数据只有同存储能力或选择的通信线路的传输率相比才有意义。为了对表 1.1 的速率有个概念，我们来看看下面的情况：目前，普通电话线路调制解调速率为 28.8kbps，在北美区域用于声音信号的传输比特率为 8kbps，这样，未经压缩的 96kbps 就需要将电话的声音信号带宽拓宽 12 倍。另外，CD-ROM 的容量大约是 650M 字节，一种正在研制的多层数字化光视频盘(DVD-5)的容量大约有 40 千兆比特，这样，对于未压缩的电视信号，CD-ROM 仅可存储 23.5 秒，DVD-5 光盘仅可存储大约 3 分钟节目。

表 1.1 未压缩信源的大致比特率

电话(200~3400Hz):	$8000 \text{ 样本数/秒} \times 12 \text{ 比特/样本} = 96\text{kbps}$
宽带语音(50~7000Hz):	$16000 \text{ 样本数/秒} \times 14 \text{ 比特/样本} = 224\text{kbps}$
宽带音频(20~20000Hz):	$44100 \text{ 样本数/秒} \times 2 \text{ 信道} \times 16 \text{ 比特/样本} = 1.412\text{Mbps}$
图像:	$512 \times 512 \text{ 像素色彩图像} \times 24 \text{ 比特/像素} = 6.3 \text{ 兆比特/图像}$
视频:	$640 \times 480 \text{ 像素色彩图像} \times 24 \text{ 比特/像素} \times 30 \text{ 图像/秒} = 221\text{Mbps}$
高清晰度电视:	$1280 \times 720 \text{ 像素色彩图像} \times 60 \text{ 图像/秒} \times 24 \text{ 比特/像素} = 1.3\text{Gbps}$

如果表 1.1 中的数据变得更大，我们将用怎样的方法来改进保持数字化传输和存储的优点呢？答案就是压缩。一般来说，压缩是信源信号，如语音、静止图像、音乐或电视的有效的数字化表示。就是说，我们用尽可能少的比特数来表示源信号并能将其还原(Berger 1971)。因此，压缩的任务就是保持信源信号在一个可以接受的状况的前提下把需要的比特数减到最少程度。这样来减少存储和传输的成本(Sayood 1996)。

你现在可能觉得压缩是必须的，接下来自然就会问：是什么原因使得压缩的应用在目前会蓬勃发展呢？答案就是：计算机，数字通信系统和电信网络，这些已出现了几十年的东西。那么，



为什么我们现在能实现这些压缩技术呢？有五个因素成就了这种可能。第一，经过四分之一一个世纪在压缩方法上出色的研究，我们终于享受到了它的成果。尽管当时没有人预见到我们今天所享用的令人惊讶的高技术产品，但是这些系统的理论研究成果却早就推导出来了，只是当时人们根本没有生产出这些产品的念头。然而如果没有牢固的基础研究，我们今天所看到的迅速发展是不可能的。第二，信号处理能力迅猛发展，从 VLSI 在今天的数字信号处理器上的应用到高速微处理器广泛应用于个人电脑，在这种计算能力下，看起来复杂的方程和算法能够不太费事地实现。第三，引入基于主观感觉的失真衡量使解决应用产品的性能问题迈出了一大步。这一点特别重要，因为大多数应用都是在声音、音乐、图像和电视方面，而最终的信息接收者是人的眼睛或耳朵。第四，建立标准减少了制造商和服务商的风险。他们在认证的产品和系统的相互操作性之中融进了压缩方法。最后一点，网络、计算机、通信方面不断地技术进步继续提供了新的机遇并提高了对数字压缩方法的接受程度。

这本书讨论压缩——各种信源的压缩，包括数据、声音、电视、静止图像和电影等等，我们将讨论基本原理，一般算法和重要的标准，这些会帮你了解目前的压缩技术和标准，以及为什么作出各种不同的设计选择。我们也希望你能改进或设计出新的压缩算法和标准。

在这一章的余下部分，我们将提出数字压缩的问题并概括地谈一下设计、实现、选择和改进数字压缩的方法、算法、标准以及服务。

## 1.2 数字压缩问题

数字压缩简单地说就是信源的有效数字表示方法。这一定义可以稍加修饰会更明确：数字压缩是在保持一个可以接受的逼真度损失的情况下采用尽可能少的比特数来表示信源。信源可以是数据、静止图像，语音、电视或其他需要存储或传输的信号。在本书中，我们用“数字压缩”这个概念既包括信源的有损压缩，也包括无损压缩——无损压缩是指能完美地恢复信源信号，有损压缩是指还原的信号与原信号不能严格一致。过去的几年里，有很多概念用作了数字压缩的同义词，有些现在仍然在用，这里我们先简要地回顾一下。

### 1.2.1 数字压缩的同义词

有两个概念现在仍经常作为数字压缩的同义词使用，就是“信号压缩”和“信号编码”（Jayant, Johnston, and Safranek 1993）。这些概念避免了与“数据”这个概念发生混淆。对一些人来说，“数据”这个词与语音、电视信号没什么关系，“信号”这个词也许有点笼统，但是这些概念已被人们接受。在一些信息文献中，“信源编码”和“逼真效果下的信源编码”是一回事，尽管在今天使用后一种说法已比 15 或 20 年前少多了，但仍有一些作者用这个概念来指代无损编码，“逼真效果下的信源编码”这个说法太长太费事，但能清楚地表明压缩是有损失的（Shannon 1959, Berger 1971, Gray 1990）。

有些研究者在语言和音频压缩方面用“信源编码”这个概念来指代编码的信源模型，这在早期的信息理论方面出现过（Flanagan 等，1979）。不过，一旦知道这种混淆，在阅读时也就好分辨了。另一些概念是无噪声和噪声编码，分别指无损和有损编码。“噪声”在这里是指解码的误差，或解码噪声。使用“噪声”一词是因为早在信源编码之前，信道的噪声编码就已经发展并为人们所接受了。有一个词是为了避免同压缩发生混淆而提出来的：“数据紧缩”，它用于表示信